$\mathcal{S}\mathcal{R}$

UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/684,488	10/04/2000	Bin Zhang	10992482-1	3131

7590   09/11/2002

HEWLETT-PACKARD COMPANY
Intellectual Property Administration
P.O. Box 272400
Fort Collins, CO   80528-9599

EXAMINER
HAMILTON, MONPLAISIR G

ART UNIT	PAPER NUMBER
2172	

DATE MAILED: 09/11/2002

Please find below and/or attached an Office communication concerning this application or proceeding.

PTO-90C (Rev. 07-01)

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM
THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on <u>04 October 2000</u> .

2a)☐ This action is **FINAL**.   2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-15* is/are pending in the application.

    4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-15* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *10/04/00* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

    Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

11)☐ The proposed drawing correction filed on _____ is: a)☐ approved b)☐ disapproved by the Examiner.

    If approved, corrected drawings are required in reply to this Office action.

12)☐ The oath or declaration is objected to by the Examiner.

Priority under 35 U.S.C. §§ 119 and 120

13)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

    a)☐ All b)☐ Some * c)☐ None of:

      1.☐ Certified copies of the priority documents have been received.

      2.☐ Certified copies of the priority documents have been received in Application No. _____ .

      3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

    * See the attached detailed Office action for a list of the certified copies not received.

14)☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application).

    a) ☐ The translation of the foreign language provisional application has been received.

15)☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121.

Attachment(s)

1)☒ Notice of References Cited (PTO-892)
2)☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3)☐ Information Disclosure Statement(s) (PTO-1449) Paper No(s) _____ .

4)☐ Interview Summary (PTO-413) Paper No(s). _____ .
5)☐ Notice of Informal Patent Application (PTO-152)
6)☐ Other: .

DETAILED ACTION

1.   Claims 1-15 are pending.

Claim Rejections - 35 USC § 103

2.   The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all

obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in
> section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are
> such that the subject matter as a whole would have been obvious at the time the invention was made to a person
> having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the
> manner in which the invention was made.

Claims 1-15 are rejected under 35 U.S.C. 103(a) as being unpatentable over US Patent

issued to Guha et al, herein referred to as Guha.

Referring to Claim 1:

Guha discloses a method that represents each cluster to be merged by a constant number

of well scattered points that capture the shape and extent of the cluster (Abstract, line 7). The

clusters with the closest pair of representative points are merged to from a new cluster (Fig 4).

This process continues until a desired number of clusters is reached (col 4, lines 40-45).

Guha does not expressly disclose the claimed "receiving a plurality of data points for

clustering; (b) receiving a size parameter for specifying the number of data points to be moved at

one time; (c) clustering the data points by using the size parameter to generate clustered results

(d) determining whether the clustered results are satisfactory;(e) when the clustered results are

satisfactory, stop clustering; (f) otherwise when the clustered results are not satisfactory revise

the size parameter. perform clustering based on the revised size parameter and the clustered

results, and proceed to step (d)."

However, the teachings disclosed by Guha are essentially the same as the claimed

limitation. The c parameter disclosed by Guha has the same functionality as the claimed size

parameter (col 6, lines 54-62).

It would have been obvious to one having ordinary skill in the art at the time that the

invention was made to modify the teachings of Guha. One of ordinary skill in the art would have

been motivated to do this because it would enable the clustering of large databases (col 1, lines

5-10).

Referring to Claim 2:

Guha discloses the limitations as discussed in Claim 1 above. Guha further discloses the

scattered points are used as representatives of the cluster. The clusters with the closest pair of

representative points are merged to reduce the number of clusters. The process continues until a

desired number of clusters is reached (col 4, lines 35-45).

Guha does not expressly disclose the claimed "(c 1) evaluating subsets of data points in

each cluster for moving into every other cluster by using a predetermined metric; wherein the

number of data points in the subset is specified by the size parameter."

However, the teachings disclosed by Guha are essentially the same as the claimed

limitation. The c parameter disclosed by Guha has the same functionality as the claimed size

parameter (col 6, lines 54-62), they both help in determining whether the functions will datasets will be merged.

It would have been obvious to one having ordinary skill in the art at the time that the invention was made to modify the teachings of Guha. One of ordinary skill in the art would have been motivated to do this because it would enable the clustering of large databases (col 1, lines 5-10).

Referring to Claim 3:

Guha discloses the limitations as discussed in Claim 2 above. Guha further discloses that the c is representative set of the cluster (col 9, lines 30-35). This set is used to calculate the distance between clusters (col 7, lines 25-37).

Guha does not expressly disclose the claimed " (c1_1) determining a geometric center of the subset of data points being evaluated for a move; (c1_2) using the geometric center of the subset of data points in the predetermined metric to generate a value."

However, the teachings disclosed by Guha are essentially the same as the claimed limitation.

It would have been obvious to one having ordinary skill in the art at the time that the invention was made to modify the teachings of Guha. One of ordinary skill in the art would have been motivated to do this because it would enable the clustering of large databases (col 1, lines 5-10).

Referring to Claim 4:

Guha discloses the limitations as discussed in Claim 3 above. Guha further discloses that a heap structure stores the distances that are calculated (col 7, lines 44-45). The method loops until only clusters remain (col 8, lines 5-10). The first element in the cluster is merged with the given cluster (col 8, lines 15-25).

Guha does not expressly disclose the claimed "(cl_ 3) determining whether the value is greater than zero; (cl_ 4) when the value is greater than zero, moving the subset of data points from a Move From cluster to a Move To cluster; (cl_5) when the value is not greater than zero, determining if there are more subsets to evaluate;(cl_6) when there are more subsets to evaluate, proceeding to step (cl); (cl_7) when there arc no more subsets to evaluate, determining whether ally point has moved; (cl_8) when a point has moved, proceeding to step (cl); and (cl_9) when no point has moved, stopping the processing."

However, the clusters as described by Guha are very similar to the claimed subsets. The functionality of the claimed limitations and that of Guha's invention are essentially the same.

It would have been obvious to one having ordinary skill in the art at the time that the invention was made to modify the teachings of Guha. One of ordinary skill in the art would have been motivated to do this because it would enable the clustering of large databases (col 1, lines 5-10).

Referring to Claim 5:

Guha discloses the limitations as discussed in Claim 4 above. Guha further discloses that that after the merge c is re-calculated (col 8, lines 10-35).

Guha does not expressly disclose the claimed "simultaneously updating the membership of at least two data points from the membership of the Move From cluster to the membership of the Move To cluster".

However, the merge operation as disclosed by Gupta is similar to the claimed move. The functionality of the claimed limitations and that of Guha's invention are essentially the same.

It would have been obvious to one having ordinary skill in the art at the time that the invention was made to modify the teachings of Guha. One of ordinary skill in the art would have been motivated to do this because it would enable the clustering of large databases (col 1, lines 5-10).

Referring to Claim 6:

Guha discloses the limitations as discussed in Claim 4 above. Guha further discloses that there is a need to compute the distance of the new cluster to all other remaining clusters. The mean is set c is also calculated for the new cluster, and all the necessary calculations and updates are carried out (col 8, lines 28-35; Fig 7).

Guha does not expressly disclose the claimed "updating the count of the Move From cluster; updating the center of the Move-From cluster; updating the count of the Move To cluster; and updating the center of the Move To cluster."

However, Guha's invention discloses limitations that are similar to claimed limitation.

It would have been obvious to one having ordinary skill in the art at the time that the invention was made to modify the teachings of Guha. One of ordinary skill in the art would have

been motivated to do this because it would enable the clustering of large databases (col 1, lines

5-10).

Referring to Claim 7:

Guha discloses the limitations as discussed in Claim 1 above. Guha further discloses that

c is shrunk toward the mean by a fraction α (col 4, lines 36-42).

Guha does not expressly disclose the claimed "decreasing the size parameter."

However, the reducing of c by α has the same functionality as the claimed limitation.

It would have been obvious to one having ordinary skill in the art at the time that the

invention was made to modify the teachings of Guha. One of ordinary skill in the art would have

been motivated to do this because it would enable the clustering of large databases (col 1, lines

5-10).

Referring to Claim 8:

Guha discloses the limitations as discussed in Claim 1 above. Guha further discloses the

metric used to determine whether or not to merge is the distance measure (col 7, lines 30-35).

This metric used the c parameter to evaluate the distance measure.

Guha does not expressly disclose the claimed "(d_1) employing a predetermined metric

for determining whether the clustered results are satisfactory; wherein the predetermined metric

includes a geometric center of the subset of points that are being evaluated for move."0

However, the limitations disclosed by Guha are essentially the same as the claimed

limitations.

It would have been obvious to one having ordinary skill in the art at the time that the invention was made to modify the teachings of Guha. One of ordinary skill in the art would have been motivated to do this because it would enable the clustering of large databases (col 1, lines 5-10).

Referring to Claim 9:

Guha discloses the limitations as discussed in Claim 8 above. Guha further discloses the expression for his metric (col 7, lines 30-35).

Guha does not expressly disclose the claimed "where U is the subset of data points being evaluated for the move, |U| is the size of U that is specified by the size parameter, m_{oo} is the geometric center of U, M_1, and m_j are the centers of the clusters and n_i, and n_j, are the counts of the clusters."

However, the metric expression disclosed by Guha is essentially the same as the claimed limitation.

It would have been obvious to one having ordinary skill in the art at the time that the invention was made to modify the teachings of Guha. One of ordinary skill in the art would have been motivated to do this because it would enable the clustering of large databases (col 1, lines 5-10).

Referring to Claim 10 and 15:

Guha discloses the limitations as discussed in Claim 1 and 11. Guha further discloses his can be used for data mining (col 1, lines 5-10).

Guha does not expressly disclose the claimed "utilized in one of a data mining

application, customer segmentation application, document categorization application, scientific

data analysis application, data compression application, vector quantization application, and

image processing application."

However, the limitations disclosed by Guha are essentially the same as the claimed

limitations.

It would have been obvious to one having ordinary skill in the art at the time that the

invention was made to modify the teachings of Guha. One of ordinary skill in the art would have

been motivated to do this because it would enable the clustering of large databases (col 1, lines

5-10).

Referring to Claim 11:

Guha discloses a method that is used for clustering a large database (col 4, lines 25-28).

The method represents each cluster to be merged by a constant number of well-scattered points

that capture the shape and extent of the cluster (Abstract, line 7). The clusters with the closest

pair of representative points are merged to from a new cluster (Fig 4). This process continues

until a desired number of clusters is reached (col 4, lines 40-45). The merge determination is

determined by a distance metric (col 7, lines 30-35) and the corresponding updates for the merge

is applied to current set of data elements (col 8, lines 15-35; Fig 7).

Guha does not expressly disclose the claimed " clustering system comprising: (a) a

source of data points to be clustered; and (b) an aggregated clustering module for moving at least

two data points at one time between a Move From cluster and a Move To cluster; wherein the

aggregated clustering module includes a move determination unit for evaluating the move of

subsets of data points from each cluster to every other cluster and determining when such a move

should be performed; and an aggregated move unit coupled to the move determination unit that

updates a Move From count, Move From center, a Move To count, and a Move To center.

However, the limitations disclosed by Guha are essentially the same as the claimed

limitations.

It would have been obvious to one having ordinary skill in the art at the time that the

invention was made to modify the teachings of Guha. One of ordinary skill in the art would have

been motivated to do this because it would enable the clustering of large databases (col 1, lines

5-10).

Referring to Claim 12:

Guha discloses the limitations as discussed in Claim 11 above. Guha further discloses

that the choice for c (representative points) is made (col 6, lines 50-61). He also teaches that

u.mean and u.rep store the mean and representative points of cluster u.

Guha does not expressly disclose the claimed "(a) a first input for receiving the data

points; (b) a second input for receiving initial center points; (c) a third input for receiving a

number of points to move at one time; (d) a parameter for storing the center point associated with

each cluster (e) a parameter for storing the count of data points associated with each cluster."

However, the limitations disclosed by Guha are essentially the same as the claimed

limitations.

It would have been obvious to one having ordinary skill in the art at the time that the

invention was made to modify the teachings of Guha. One of ordinary skill in the art would have

been motivated to do this because it would enable the clustering of large databases (col 1, lines

5-10).

Referring to Claim 13:

Guha discloses the limitations as discussed in Claim 11 above. Guha further discloses

that the distance metric uses the c representative points to make the merge determination (col 7,

lines 25-40).

Guha does not expressly disclose the claimed "a geometric center determination unit for

determining the geometric center of a current subset of data points and providing the

geometric center to the move determination unit for move evaluation and move

determination."

However, the limitations disclosed by Guha are essentially the same as the claimed

limitations.

It would have been obvious to one having ordinary skill in the art at the time that the

invention was made to modify the teachings of Guha. One of ordinary skill in the art would have

been motivated to do this because it would enable the clustering of large databases (col 1, lines

5-10).

Referring to Claim 14:

Guha discloses the limitations as discussed in Claim 13 above. Guha further discloses
the metric used to determine whether or not to merge is the distance measure (col 7, lines 30-35).
This metric used the c parameter to evaluate the distance measure.

Guha does not expressly disclose the claimed "a move evaluation mechanism for
employing a predetermined metric for move evaluation; wherein the predetermined metric
includes the geometric center of a current subset of data points."

However, the limitations disclosed by Guha are essentially the same as the claimed
limitations.

It would have been obvious to one having ordinary skill in the art at the time that the
invention was made to modify the teachings of Guha. One of ordinary skill in the art would have
been motivated to do this because it would enable the clustering of large databases (col 1, lines
5-10).

Prior Art

3.   The prior art made of record and not relied upon is considered pertinent to applicant's
disclosure.

Wakahara et al., Extended Mean Shift in Handwriting Clustering. Wakahara discloses a
window parameter controlling the size of the neighborhood.

MacQueen J. Some methods for Classification and Analysis of Multivariate
Observations. MacQueen discloses starting with an arbitrary partition into k sets. The means of
the points in each set are computer, a new partition of the points is formed by the rule of putting
the points into groups on the basis of nearness to the set of first means. [Compute the means of

the points in each set in the initial partition and reclassify the points on the basis on nearness.]
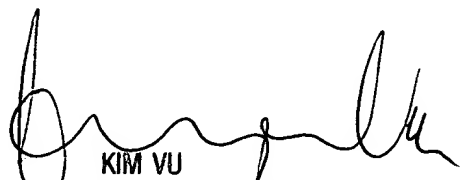
(Section 3.6)

Conclusion

4.   Any inquiry concerning this communication or earlier communications from the

examiner should be directed to Monplaisir G Hamilton whose telephone number is 1703-305-

5116. The examiner can normally be reached on Monday - Friday (8:00 am - 4:30 pm).

If attempts to reach the examiner by telephone are unsuccessful, the examiner's

supervisor, Kim Y Vu can be reached on 1703-305-4393. The fax phone numbers for the

organization where this application or proceeding is assigned are 1703-746-7239 for regular

communications and 1703-746-7238 for After Final communications.

Any inquiry of a general nature or relating to the status of this application or proceeding

should be directed to the receptionist whose telephone number is 1703-305-3900.

Monplaisir Hamilton
September 6, 2002

KIM VU
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2100